

SEQUENCE LISTING

<110> Messier, Walter
Sikela, James M

<120> Methods to Identify Polynucleotide and Polypeptide Sequences Which May Be Associated with Physiological and Medical Conditions

<130> GENO 200.2/CIP

<140>
<141>

<150> 09/591,435
<151> 2000-06-09

<150> 09/240,915
<151> 1999-01-29

<150> 60/073,263
<151> 1998-01-30

<150> 60/098,987
<151> 1998-09-02

<160> 30

<170> PatentIn Ver. 2.0

<210> 1
<211> 1518
<212> DNA
<213> Homo sapiens

<400> 1
cagacatctg tgtccccctc aaaagtcatc ctgccccggg gaggctccgt gctggtgaca 60
tgcagcacct cctgtgacca gcccaagttg ttgggcatacg agaccccgtt gcctaaaaag 120
gagttgctcc tgcctggaa caaccggaag gtgtatgaac tgagcaatgt gcaagaagat 180
agccaaccaa tgtgctattc aaactgccct gatggcagt caacagctaa aacttcctc 240
accgtgtact ggactccaga acgggtggaa ctggcacccc tcccctttg gcagccagtg 300
ggcaagaacc ttaccctacg ctgccaggtg gagggtgggg caccggggc caacccacc 360

gtggtgctgc tccgtgggga gaaggagctg aaacgggagc cagctgtggg ggagcccgt 420
gaggtcacga ccacggtgct ggtgaggaga gatcaccatg gagccaattt ctctgtccgc 480
actgaactgg acctgcggcc ccaagggctg gagctgttg agaacacacctc ggccccctac 540
cagctccaga cctttgtcct gccagcgact cccccacaac ttgtcagccc ccgggtccta 600
gaggtggaca cgcaaaaaac cgtggctgt tccctggacg ggctgttccc agtctcgag 660
gcccagggtcc acctggact gggggaccag aggtaacc ccacagtcac ctatggcaac 720
gactccttct cggccaaggc ctcagtcagt gtgaccgcag aggacgaggg cacccagcgg 780
ctgacgtgtg cagtaatact gggaaaccag agccaggaga cactgcagac agtgaccatc 840
tacagcttcc cggcgccaa cgtgattctg acgaagccag aggtctcaga agggaccgag 900
gtgacagtga agtgtgaggc ccaccctaga gccaagggtga cgctgaatgg gttccagcc 960
cagccactgg gcccggggc ccagctcctg ctgaaggcca ccccagagga caacggcgc 1020
agcttctcct gctctgcaac cctggaggtg gccggccagc ttatacacaa gaaccagacc 1080
cgggagcttc gtgtcctgta tgccccca ctggacgaga gggattgtcc gggaaactgg 1140
acgtggccag aaaattccca gcagactcca atgtgccagg cttgggggaa cccattgccc 1200
gagctcaagt gtctaaagga tggactttc ccactgccc tcggggaaatc agtgaactgtc 1260
actcgagatc ttgagggcac ctacctctgt cggccagga gcactcaagg ggaggtcacc 1320
cgcgaggtga ccgtaatgt gctctcccc cggtatgaga ttgtcatcat cactgtggta 1380
cgagccgcag tcataatggg cactgcaggc ctcagcacgt acctctataa ccgccagcgg 1440
aagatcaaga aatacagact acaacaggcc caaaaaggga cccccatgaa accgaacaca 1500
caagccacgc ctccctga 1518

<210> 2
<211> 1518
<212> DNA
<213> Pan troglodytes

<220>
<221> CDS

<222> (1)..(1518)

<400> 2

cag aca tct gtg tcc ccc cca aaa gtc atc ctg ccc cg^g gga ggc tcc 48
Gln Thr Ser Val Ser Pro Pro Lys Val Ile Leu Pro Arg Gly Gly Ser
1 5 10 15

gtg cag gtg aca tgc agc acc tcc tgt gac cag ccc gac ttg ttg ggc 96
Val Gln Val Thr Cys Ser Thr Ser Cys Asp Gln Pro Asp Leu Leu Gly
20 25 30

ata gag acc ccg ttg cct aaa aag gag ttg ctt ctg ggt ggg aac aac 144
Ile Glu Thr Pro Leu Pro Lys Lys Glu Leu Leu Gly Gly Asn Asn
35 40 45

tgg aag gtg tat gaa ctg agc aat gtg caa gaa gat agc caa cca atg 192
Trp Lys Val Tyr Glu Leu Ser Asn Val Gln Glu Asp Ser Gln Pro Met
50 55 60

tgc tat tca aac tgc cct gat ggg cag tca aca gct aaa acc ttc ctc 240
Cys Tyr Ser Asn Cys Pro Asp Gly Gln Ser Thr Ala Lys Thr Phe Leu
65 70 75 80

acc gtg tac tgg act cca gaa cgg gtg gaa ctg gca ccc ctc ccc tct 288
Thr Val Tyr Trp Thr Pro Glu Arg Val Glu Leu Ala Pro Leu Pro Ser
85 90 95

tgg cag cca gtg ggc aag gac ctt acc cta cgc tgc cag gtg gag ggt 336
Trp Gln Pro Val Gly Lys Asp Leu Thr Leu Arg Cys Gln Val Glu Gly
100 105 110

ggg gca ccc cgg gcc aac ctc acc gtg gtg ctc cgt ggg gag aag 384
Gly Ala Pro Arg Ala Asn Leu Thr Val Val Leu Leu Arg Gly Glu Lys
115 120 125

gag ctg aaa cgg gag cca gct gtg ggg gag ccc gct gag gtc acg acc 432
Glu Leu Lys Arg Glu Pro Ala Val Gly Glu Pro Ala Glu Val Thr Thr
130 135 140

acg gtg ctg gtg gag aga gat cac cat gga gcc aat ttc tcg tgc cgc 480
Thr Val Leu Val Glu Arg Asp His His Gly Ala Asn Phe Ser Cys Arg
145 150 155 160

act gaa ctg gac ctg cgg ccc caa ggg ctg cag ctg ttt gag aac acc 528
Thr Glu Leu Asp Leu Arg Pro Gln Gly Leu Gln Leu Phe Glu Asn Thr
165 170 175

tcg gcc ccc cac cag ctc caa acc ttt gtc ctg cca gc^g act ccc cca 576

Ser Ala Pro His Gln Leu Gln Thr Phe Val Leu Pro Ala Thr Pro Pro
 180 185 190
 caa ctt gtc agc ccc cggtt cta gag gtg gac acg cag ggg acc gtc 624
 Gln Leu Val Ser Pro Arg Val Leu Glu Val Asp Thr Gln Gly Thr Val
 195 200 205
 gtc tgt tcc ctg gac ggg ctg ttc cca gtc tcg gag gcc cag gtc cac 672
 Val Cys Ser Leu Asp Gly Leu Phe Pro Val Ser Glu Ala Gln Val His
 210 215 220
 ctg gca ctg ggg gac cag agg ttg aac ccc aca gtc acc tat ggc aat 720
 Leu Ala Leu Gly Asp Gln Arg Leu Asn Pro Thr Val Thr Tyr Gly Asn
 225 230 235 240
 gac tcc ttc tcg gcc aag gcc tca gtc agt gtg acc gca gag gac gag 768
 Asp Ser Phe Ser Ala Lys Ala Ser Val Ser Val Thr Ala Glu Asp Glu
 245 250 255
 ggc acc cag cgg ctg acg tgt gca gta ata ctg ggg aac cag agc cgg 816
 Gly Thr Gln Arg Leu Thr Cys Ala Val Ile Leu Gly Asn Gln Ser Arg
 260 265 270
 gag aca ctg cag aca gtg acc atc tac agc ttt ccg gcg ccc aac gtg 864
 Glu Thr Leu Gln Thr Val Thr Ile Tyr Ser Phe Pro Ala Pro Asn Val
 275 280 285
 att ctg acg aag cca gag gtc tca gaa ggg acc gag gtg aca gtg aag 912
 Ile Leu Thr Lys Pro Glu Val Ser Glu Gly Thr Glu Val Thr Val Lys
 290 295 300
 tgt gag gcc cac cct aga gcc aag gtg acg ctg aat ggg gtt cca gcc 960
 Cys Glu Ala His Pro Arg Ala Lys Val Thr Leu Asn Gly Val Pro Ala
 305 310 315 320
 cag cca gtg ggc ccg agg gtc cag ctc ctg ctg aag gcc acc cca gag 1008
 Gln Pro Val Gly Pro Arg Val Gln Leu Leu Lys Ala Thr Pro Glu
 325 330 335
 gac aac ggg cgc agc ttc tcc tgc tct gca acc ctg gag gtg gcc ggc 1056
 Asp Asn Gly Arg Ser Phe Ser Cys Ser Ala Thr Leu Glu Val Ala Gly
 340 345 350
 cag ctt ata cac aag aac cag acc cgg gag ctt cgt gtc ctg tat ggc 1104
 Gln Leu Ile His Lys Asn Gln Thr Arg Glu Leu Arg Val Leu Tyr Gly
 355 360 365
 ccc cga ctg gac gag agg gat tgt ccg gga aac tgg acg tgg cca gaa 1152

D G S H T R E D E E D E E D E E D E E D E E

Pro Arg Leu Asp Glu Arg Asp Cys Pro Gly Asn Trp Thr Trp Pro Glu				
370	375	380		
aat tcc cag cag act cca atg tgc cag gct tcg ggg aac cca ttg ccc				1200
Asn Ser Gln Gln Thr Pro Met Cys Gln Ala Ser Gly Asn Pro Leu Pro				
385	390	395	400	
gag ctc aag tgt cta aag gat ggc act ttc cca ctg ccc gtc ggg gaa				1248
Glu Leu Lys Cys Leu Lys Asp Gly Thr Phe Pro Leu Pro Val Gly Glu				
405	410	415		
tca gtg act gtc act cga gat ctt gag ggc acc tac ctc tgt cgg gcc				1296
Ser Val Thr Val Thr Arg Asp Leu Glu Gly Thr Tyr Leu Cys Arg Ala				
420	425	430		
agg agc actcaa ggg gag gtc acc cgc aag gtg acc gtg aat gtg ctc				1344
Arg Ser Thr Gln Gly Glu Val Thr Arg Lys Val Thr Val Asn Val Leu				
435	440	445		
tcc ccc cgg tat gag att gtc atc atc act gtg gta gca gcc gca gtc				1392
Ser Pro Arg Tyr Glu Ile Val Ile Ile Thr Val Val Ala Ala Ala Val				
450	455	460		
ata atg ggc act gca ggc ctc agc acg tac ctc tat aac cgc cag cgg				1440
Ile Met Gly Thr Ala Gly Leu Ser Thr Tyr Leu Tyr Asn Arg Gln Arg				
465	470	475	480	
aag atc agg aaa tac aga cta caa cag gct caa aaa ggg acc ccc atg				1488
Lys Ile Arg Lys Tyr Arg Leu Gln Gln Ala Gln Lys Gly Thr Pro Met				
485	490	495		
aaa ccg aac aca caa gcc acg cct ccc tga				1518
Lys Pro Asn Thr Gln Ala Thr Pro Pro				
500	505			

<210> 3

<211> 505

<212> PRT

<213> Pan troglodytes

<400> 3

Gln Thr Ser Val Ser Pro Pro Lys Val Ile Leu Pro Arg Gly Gly Ser			
1	5	10	15

Val Gln Val Thr Cys Ser Thr Ser Cys Asp Gln Pro Asp Leu Leu Gly			
20	25	30	

Ile Glu Thr Pro Leu Pro Lys Lys Glu Leu Leu Leu Gly Gly Asn Asn
35 40 45

Trp Lys Val Tyr Glu Leu Ser Asn Val Gln Glu Asp Ser Gln Pro Met
50 55 60

Cys Tyr Ser Asn Cys Pro Asp Gly Gln Ser Thr Ala Lys Thr Phe Leu
65 70 75 80

Thr Val Tyr Trp Thr Pro Glu Arg Val Glu Leu Ala Pro Leu Pro Ser
85 90 95

Trp Gln Pro Val Gly Lys Asp Leu Thr Leu Arg Cys Gln Val Glu Gly
100 105 110

Gly Ala Pro Arg Ala Asn Leu Thr Val Val Leu Leu Arg Gly Glu Lys
115 120 125

Glu Leu Lys Arg Glu Pro Ala Val Gly Glu Pro Ala Glu Val Thr Thr
130 135 140

Thr Val Leu Val Glu Arg Asp His His Gly Ala Asn Phe Ser Cys Arg
145 150 155 160

Thr Glu Leu Asp Leu Arg Pro Gln Gly Leu Gln Leu Phe Glu Asn Thr
165 170 175

Ser Ala Pro His Gln Leu Gln Thr Phe Val Leu Pro Ala Thr Pro Pro
180 185 190

Gln Leu Val Ser Pro Arg Val Leu Glu Val Asp Thr Gln Gly Thr Val
195 200 205

Val Cys Ser Leu Asp Gly Leu Phe Pro Val Ser Glu Ala Gln Val His
210 215 220

Leu Ala Leu Gly Asp Gln Arg Leu Asn Pro Thr Val Thr Tyr Gly Asn
225 230 235 240

Asp Ser Phe Ser Ala Lys Ala Ser Val Ser Val Thr Ala Glu Asp Glu
245 250 255

Gly Thr Gln Arg Leu Thr Cys Ala Val Ile Leu Gly Asn Gln Ser Arg
260 265 270

Glu Thr Leu Gln Thr Val Thr Ile Tyr Ser Phe Pro Ala Pro Asn Val
275 280 285

Ile Leu Thr Lys Pro Glu Val Ser Glu Gly Thr Glu Val Thr Val Lys
 290 295 300

 Cys Glu Ala His Pro Arg Ala Lys Val Thr Leu Asn Gly Val Pro Ala
 305 310 315 320

 Gln Pro Val Gly Pro Arg Val Gln Leu Leu Leu Lys Ala Thr Pro Glu
 325 330 335

 Asp Asn Gly Arg Ser Phe Ser Cys Ser Ala Thr Leu Glu Val Ala Gly
 340 345 350

 Gln Leu Ile His Lys Asn Gln Thr Arg Glu Leu Arg Val Leu Tyr Gly
 355 360 365

 Pro Arg Leu Asp Glu Arg Asp Cys Pro Gly Asn Trp Thr Trp Pro Glu
 370 375 380

 Asn Ser Gln Gln Thr Pro Met Cys Gln Ala Ser Gly Asn Pro Leu Pro
 385 390 395 400

 Glu Leu Lys Cys Leu Lys Asp Gly Thr Phe Pro Leu Pro Val Gly Glu
 405 410 415

 Ser Val Thr Val Thr Arg Asp Leu Glu Gly Thr Tyr Leu Cys Arg Ala
 420 425 430

 Arg Ser Thr Gln Gly Glu Val Thr Arg Lys Val Thr Val Asn Val Leu
 435 440 445

 Ser Pro Arg Tyr Glu Ile Val Ile Ile Thr Val Val Ala Ala Ala Val
 450 455 460

 Ile Met Gly Thr Ala Gly Leu Ser Thr Tyr Leu Tyr Asn Arg Gln Arg
 465 470 475 480

 Lys Ile Arg Lys Tyr Arg Leu Gln Gln Ala Gln Lys Gly Thr Pro Met
 485 490 495

 Lys Pro Asn Thr Gln Ala Thr Pro Pro
 500 505

<210> 4
 <211> 1515
 <212> DNA
 <213> Gorilla gorilla

<400> 4
cagacatctg tgcgtccccc aaaagtcatc ctgccccggg gaggctccgt gctggtgaca 60
tgcagcacct cctgtgacca gcccaccttg ttgggcatacg agaccccggtt gcctaaaaag 120
gagttgctcc tgcttggaa caaccagaag gtgtatgaac tgagcaatgt gcaagaagat 180
agccaaccaa tgtgttattc aaactgcctt gatggcagt caacagctaa aaccccttc 240
accgtgtact ggactccaga acgggtggaa ctggcaccccc tccctcttg gcagccagt 300
ggcaaggacc ttaccctacg ctgccaggtg gagggtgggg caccccgccc caacctcatc 360
gtggtgctgc tccgtggggaa ggaggagctg aaacgggagc cagctgtggg ggagccgc 420
gaggtcacga ccacggtgcc ggtggagaaa gatcaccatg gagccaattt cttgtgccgc 480
actgaactgg acctgcggcc ccaagggctg aagctgtttg agaacacctc ggccccctac 540
cagctccaaa cctttgtcct gccagcgact ccccccacaac ttgtcagccc tcgggtccta 600
gaggtggaca cgcaggggac tgtggtctgt tccctggacg ggctgttccc agtctcgag 660
gcccagggtcc acctggact gggggaccag aggttgaacc ccacagtcac ctatggcaac 720
gactccttct cagccaaaggc ctcagtcagt gtgaccgcag aggacgaggg caccctgtgg 780
ctgacgtgtg cagtaatact ggggacccag agccaggaga cactgcagac agtgaccatc 840
tacagctttc cggcacccaa cgtgattctg acgaagccag aggtctcaga agggaccgag 900
gtgacagtga agtgtgaggc ccaccctaga gccaaaggta cactgaatgg gttccagcc 960
cagccacccgg gcccggggac ccagttcctg ctgaaggcca cccagagga caacgggcgc 1020
agcttctcct gctctgcaac cctggagggtg gcccggccagc ttatacacaa gaaccagacc 1080
cgggagcttc gtgtcctgtt tggccccccga ctggatgaga gggattgtcc gggaaactgg 1140
acgtggccag aaaattccca gcagactcca atgtgccagg cttggggggaa cccattgccc 1200
gagctcaagt gtctaaagga tggcactttc ccactgcccc tcggggaaatc agtactgtc 1260
actcgagatc ttgagggcac ctacctctgt cggccagga gcactcaagg ggaggtcacc 1320
cgcgaggtga ccgtgaatgt gctctccccc cggtatgagt ttgtcatcat cgctgtggta 1380
gcagccgcag tcataatggg cactgcaggc ctcagcacgt acctctataa ccggccagcgg 1440

aagatcgga aatacagact acaacaggct caaaaagggaa ccccatgaa accgaacaca 1500
caagccacgc ctccc 1515

<210> 5
<211> 1515
<212> DNA
<213> Pongo pygmaeus

<400> 5
cacacatctg tgtcctccgc caacgtcttc ctgccccggg gaggctccgt gctagtgaat 60
tgccagcacct cctgtgacca gcccaccttg ttgggcatacg agaccccgtt gcctaaaaag 120
gagttgtcc cggttggaa caactggaag atgtatgaac tgagcaatgt gcaagaagat 180
agccaaccaa tgtgctattc aaactgccct gatggcagt cagcagctaa aacccctc 240
accgtgtact ggactccaga acgggtggaa ctggcaccccc tcccccttttgcagccagtg 300
ggcaagaacc ttaccctacg ctgccaggtg gaggggtgggg caccccgggc caacccctcacc 360
gtggtattgc tccgtggggaa ggaggagctg agccggcagc cagcggtggg ggagcccgcc 420
gaggtcacgg ccacggtgct ggcgaggaaa gatgaccacg gagccaattt ctcgtgccgc 480
actgaactgg acctgcggcc ccaagggctg gagctgttttga aacacacccggccccac 540
cagctccaaa ctttgcct gccagcgact ccccaacaac ttgtcagccc cgggtccta 600
gaggtggaca cgcaggggac cgtggctgt tccctggacg ggctgttccc agtctcgag 660
gcccagggtcc acttggcact gggggaccag aggttgaacc ccacagtcac ctatggcgac 720
gactccctct cggccaaggc ctcagtcagt gtgaccgcag aggaggaggg cacccagtgg 780
ctgtggtgtg cagtatact gaggaaccag agccaggaga cacggcagac agtaccatc 840
tacagcttc ctgcacccaa cgtgactctg atgaagccag aggtctcaga agggaccgag 900
gtgatagtga agtgtgagggc ccaccctgca gccaacgtga cgctgaatgg ggttccagcc 960
cagccgcccgg gcccggggc ccagttcctg ctgaaggcca ccccaaggaga caacgggcgc 1020
agtttcctct gctctgcaac cctggaggtg gccggccagc ttatacaca gaaccagacc 1080

cgggagcttc gagtcctgta tggcccccga ctggacgaga gggattgtcc gggaaactgg 1140
acgtggccag aaaactccca gcagactcca atgtgccagg cttggggaa ccccttgccc 1200
gagctcaagt gtctaaagga tggcactttc ccactgccc tcgggaaatc agtactgtc 1260
actcgagatc ttgagggcac ctacctctgt cggccagga gcactcaagg ggaggtcacc 1320
cgcgaggtga ccgtaatgt gctctccccc cggtatgaga ttgtcatcat cactgtggta 1380
gcagccgcag ccatactggg cactgcagge ctcagcacgt acctctataa ccgccagcgg 1440
aagatcagga tatacagact acaaaggct caaaaaggaa cccccatgaa accaaacaca 1500
caaaccacgc ctccc 1515

<210> 6
<211> 505
<212> PRT
<213> Homo sapiens

<400> 6
Gln Thr Ser Val Ser Pro Ser Lys Val Ile Leu Pro Arg Gly Gly Ser
1 5 10 15

Val Leu Val Thr Cys Ser Thr Ser Cys Asp Gln Pro Lys Leu Leu Gly
20 25 30

Ile Glu Thr Pro Leu Pro Lys Lys Glu Leu Leu Leu Pro Gly Asn Asn
35 40 45

Arg Lys Val Tyr Glu Leu Ser Asn Val Gln Glu Asp Ser Gln Pro Met
50 55 60

Cys Tyr Ser Asn Cys Pro Asp Gly Gln Ser Thr Ala Lys Thr Phe Leu
65 70 75 80

Thr Val Tyr Trp Thr Pro Glu Arg Val Glu Leu Ala Pro Leu Pro Ser
85 90 95

Trp Gln Pro Val Gly Lys Asn Leu Thr Leu Arg Cys Gln Val Glu Gly
100 105 110

Gly Ala Pro Arg Ala Asn Leu Thr Val Val Leu Leu Arg Gly Glu Lys
115 120 125

Glu Leu Lys Arg Glu Pro Ala Val Gly Glu Pro Ala Glu Val Thr Thr

130 135 140

Thr Val Leu Val Arg Arg Asp His His Gly Ala Asn Phe Ser Cys Arg
145 150 155 160

Thr Glu Leu Asp Leu Arg Pro Gln Gly Leu Glu Leu Phe Glu Asn Thr
165 170 175

Ser Ala Pro Tyr Gln Leu Gln Thr Phe Val Leu Pro Ala Thr Pro Pro
180 185 190

Gln Leu Val Ser Pro Arg Val Leu Glu Val Asp Thr Gln Gly Thr Val
195 200 205

Val Cys Ser Leu Asp Gly Leu Phe Pro Val Ser Glu Ala Gln Val His
210 215 220

Leu Ala Leu Gly Asp Gln Arg Leu Asn Pro Thr Val Thr Tyr Gly Asn
225 230 235 240

Asp Ser Phe Ser Ala Lys Ala Ser Val Ser Val Thr Ala Glu Asp Glu
245 250 255

Gly Thr Gln Arg Leu Thr Cys Ala Val Ile Leu Gly Asn Gln Ser Gln
260 265 270

Glu Thr Leu Gln Thr Val Thr Ile Tyr Ser Phe Pro Ala Pro Asn Val
275 280 285

Ile Leu Thr Lys Pro Glu Val Ser Glu Gly Thr Glu Val Thr Val Lys
290 295 300

Cys Glu Ala His Pro Arg Ala Lys Val Thr Leu Asn Gly Val Pro Ala
305 310 315 320

Gln Pro Leu Gly Pro Arg Ala Gln Leu Leu Lys Ala Thr Pro Glu
325 330 335

Asp Asn Gly Arg Ser Phe Ser Cys Ser Ala Thr Leu Glu Val Ala Gly
340 345 350

Gln Leu Ile His Lys Asn Gln Thr Arg Glu Leu Arg Val Leu Tyr Gly
355 360 365

Pro Arg Leu Asp Glu Arg Asp Cys Pro Gly Asn Trp Thr Trp Pro Glu
370 375 380

Asn Ser Gln Gln Thr Pro Met Cys Gln Ala Trp Gly Asn Pro Leu Pro

385 390 395 400

Glu Leu Lys Cys Leu Lys Asp Gly Thr Phe Pro Leu Pro Ile Gly Glu
405 410 415

Ser Val Thr Val Thr Arg Asp Leu Glu Gly Thr Tyr Leu Cys Arg Ala
420 425 430

Arg Ser Thr Gln Gly Glu Val Thr Arg Glu Val Thr Val Asn Val Leu
435 440 445

Ser Pro Arg Tyr Glu Ile Val Ile Ile Thr Val Val Ala Ala Ala Val
450 455 460

Ile Met Gly Thr Ala Gly Leu Ser Thr Tyr Leu Tyr Asn Arg Gln Arg
465 470 475 480

Lys Ile Lys Lys Tyr Arg Leu Gln Gln Ala Gln Lys Gly Thr Pro Met
485 490 495

Lys Pro Asn Thr Gln Ala Thr Pro Pro
500 505

<210> 7
<211> 254
<212> PRT
<213> Homo sapiens

<400> 7

Ser Asp Glu Lys Val Phe Glu Val His Val Arg Pro Lys Lys Leu Ala
1 5 10 15

Val Glu Pro Lys Gly Ser Leu Glu Val Asn Cys Ser Thr Thr Cys Asn
20 25 30

Gln Pro Glu Val Gly Gly Leu Glu Thr Ser Leu Asp Lys Ile Leu Leu
35 40 45

Asp Glu Gln Ala Gln Trp Lys His Tyr Leu Val Ser Asn Ile Ser His
50 55 60

Asp Thr Val Leu Gln Cys His Phe Thr Cys Ser Gly Lys Gln Glu Ser
65 70 75 80

Met Asn Ser Asn Val Ser Val Tyr Gln Pro Pro Arg Gln Val Ile Leu
85 90 95

Thr Leu Gln Pro Thr Leu Val Ala Val Gly Lys Ser Phe Thr Ile Glu
100 105 110

Cys Arg Val Pro Thr Val Glu Pro Leu Asp Ser Leu Thr Leu Phe Leu
115 120 125

Phe Arg Gly Asn Glu Thr Leu His Tyr Glu Thr Phe Gly Lys Ala Ala
130 135 140

Pro Ala Pro Gln Glu Ala Thr Ala Thr Phe Asn Ser Thr Ala Asp Arg
145 150 155 160

Glu Asp Gly His Arg Asn Phe Ser Cys Leu Ala Val Leu Asp Leu Met
165 170 175

Ser Arg Gly Gly Asn Ile Phe His Lys His Ser Ala Pro Lys Met Leu
180 185 190

Glu Ile Tyr Glu Pro Val Ser Asp Ser Gln Met Val Ile Ile Val Thr
195 200 205

Val Val Ser Val Leu Leu Ser Leu Phe Val Thr Ser Val Leu Leu Cys
210 215 220

Phe Ile Phe Gly Gln His Leu Arg Gln Gln Arg Met Gly Thr Tyr Gly
225 230 235 240

Val Arg Ala Ala Trp Arg Arg Leu Pro Gln Ala Phe Arg Pro
245 250

<210> 8
<211> 518
<212> PRT
<213> Homo sapiens

<400> 8
Gln Glu Phe Leu Leu Arg Val Glu Pro Gln Asn Pro Val Leu Ser Ala
1 5 10 15

Gly Gly Ser Leu Phe Val Asn Cys Ser Thr Asp Cys Pro Ser Ser Glu
20 25 30

Lys Ile Ala Leu Glu Thr Ser Leu Ser Lys Glu Leu Val Ala Ser Gly
35 40 45

Met Gly Trp Ala Ala Phe Asn Leu Ser Asn Val Thr Gly Asn Ser Arg
50 55 60

Ile Leu Cys Ser Val Tyr Cys Asn Gly Ser Gln Ile Thr Gly Ser Ser
 65 70 75 80

Asn Ile Thr Val Tyr Gly Leu Pro Glu Arg Val Glu Leu Ala Pro Leu
 85 90 95

Pro Pro Trp Gln Pro Val Gly Gln Asn Phe Thr Leu Arg Cys Gln Val
 100 105 110

Glu Gly Gly Ser Pro Arg Thr Ser Leu Thr Val Val Leu Leu Arg Trp
 115 120 125

Glu Glu Glu Leu Ser Arg Gln Pro Ala Val Glu Glu Pro Ala Glu Val
 130 135 140

Thr Ala Thr Val Leu Ala Ser Arg Asp Asp His Gly Ala Pro Phe Ser
 145 150 155 160

Cys Arg Thr Glu Leu Asp Met Gln Pro Gln Gly Leu Gly Leu Phe Val
 165 170 175

Asn Thr Ser Ala Pro Arg Gln Leu Arg Thr Phe Val Leu Pro Val Thr
 180 185 190

Pro Pro Arg Leu Val Ala Pro Arg Phe Leu Glu Val Glu Thr Ser Trp
 195 200 205

Pro Val Asp Cys Thr Leu Asp Gly Leu Phe Pro Ala Ser Glu Ala Gln
 210 215 220

Val Tyr Leu Ala Leu Gly Asp Gln Met Leu Asn Ala Thr Val Met Asn
 225 230 235 240

His Gly Asp Thr Leu Thr Ala Thr Ala Thr Ala Arg Ala Asp
 245 250 255

Gln Glu Gly Ala Arg Glu Ile Val Cys Asn Val Thr Leu Gly Gly Glu
 260 265 270

Arg Arg Glu Ala Arg Glu Asn Leu Thr Val Phe Ser Phe Leu Gly Pro
 275 280 285

Ile Val Asn Leu Ser Glu Pro Thr Ala His Glu Gly Ser Thr Val Thr
 290 295 300

Val Ser Cys Met Ala Gly Ala Arg Val Gln Val Thr Leu Asp Gly Val
 305 310 315 320

Pro Ala Ala Ala Pro Gly Gln Pro Ala Gln Leu Gln Leu Asn Ala Thr
325 330 335

Glu Ser Asp Asp Gly Arg Ser Phe Phe Cys Ser Ala Thr Leu Glu Val
340 345 350

Asp Gly Glu Phe Leu His Arg Asn Ser Ser Val Gln Leu Arg Val Leu
355 360 365

Tyr Gly Pro Lys Ile Asp Arg Ala Thr Cys Pro Gln His Leu Lys Trp
370 375 380

Lys Asp Lys Thr Arg His Val Leu Gln Cys Gln Ala Arg Gly Asn Pro
385 390 395 400

Tyr Pro Glu Leu Arg Cys Leu Lys Glu Gly Ser Ser Arg Glu Val Pro
405 410 415

Val Gly Ile Pro Phe Phe Val Asn Val Thr His Asn Gly Thr Tyr Gln
420 425 430

Cys Gln Ala Ser Ser Ser Arg Gly Lys Tyr Thr Leu Val Val Val Met
435 440 445

Asp Ile Glu Ala Gly Ser Ser His Phe Val Pro Val Phe Val Ala Val
450 455 460

Leu Leu Thr Leu Gly Val Val Thr Ile Val Leu Ala Leu Met Tyr Val
465 470 475 480

Phe Arg Glu His Gln Arg Ser Gly Ser Tyr His Val Arg Glu Glu Ser
485 490 495

Thr Tyr Leu Pro Leu Thr Ser Met Gln Pro Thr Glu Ala Met Gly Glu
500 505 510

Glu Pro Ser Arg Ala Glu
515

<210> 9
<211> 1212
<212> DNA
<213> Homo sapiens

<400> 9
atgagtgact ccaaggaacc aagactgcag cagctgggcc tcctggagga ggaacagctg 60

agaggcctg gattccgaca gactcgagga tacaagagct tagcagggtg tcttggccat 120
ggtccccctgg tgctgcaact cctctccttc acgctcttgg ctgggctcct tgtccaagtg 180
tccaagggtcc ccagctccat aagtcaggaa caatccaggc aagacgcgat ctaccagaac 240
ctgacccagc ttAAAGCTGC agtgggttag ctctcagaga aatccaagct gcaggagatc 300
taccaggagc tgacccagct gaaggctgca gtgggttagc ttccagagaa atctaagctg 360
caggagatct accaggagct gacccggctg aaggctgcag tgggttagct tccagagaaa 420
tctaaGCTGC aggagatcta ccaggagctg acctggctga aggctgcagt gggtagctt 480
ccagagaaat ctaagatgca ggagatctac caggagctga ctcggctgaa ggctgcagtg 540
ggtagcttc cagagaaatc taagcagcag gagatctacc aggagctgac ccggctgaag 600
gctgcagtgg gttagcttc agagaaatct aagcagcagg agatctacca ggagctgacc 660
cggtgaagg ctgcagtgg tgagcttcca gagaaatcta agcagcagga gatctaccag 720
gagctgaccc agctgaaggc tgcagtggaa cgccctgtgcc accccctgtcc ctgggaatgg 780
acattcttcc aaggaaactg ttacttcatg tctaactccc agcggaaactg gcacgactcc 840
atcacccgct gcaaagaagt gggggcccaag ctcgtcgtaa tcaaaagtgc tgaggagcag 900
aacttcc tac agctgcagtc ttccagaagt aaccgcttca cctggatggg actttcagat 960
ctaaatcagg aaggcacgtg gcaatgggtg gacggctcac ctctgttgcc cagcttcaag 1020
cagtatttgg aacagaggaga gcccaacaac gttggggagg aagactgcgc ggaatttagt 1080
ggcaatggct ggaacgacga caaatgtaat cttgccaaat tctggatctg caaaaagtcc 1140
gcagcctcct gctccaggaa tgaagaacag tttcttctc cagccctgc caccccaaac 1200
ccccctcctg cg 1212

<210> 10
<211> 1212
<212> DNA
<213> Pan troglodytes

<400> 10

YODA - DRAFT

atgagtgact ccaaggaacc aagactgcag cagctgggcc tcctggagga ggaacagctg 60
agaggccttg gattccgaca gactcgaggc tacaagagct tagcaggggtg tcttggccat 120
ggtccccctgg tgctgcaact cctctccttc acgctttgg ctgggctcct tgtccaagtg 180
tccaagggtcc ccagctccat aagtcagggaa gaatccaggc aagacgtgat ctaccagaac 240
ctgaccaggc tttaaagctgc agtgggtgag ctctcagaga aatccaagct gcaggagatc 300
taccaggagc tgacccagct gaaggctgca gtgggtgagc ttccagagaa atctaaggcag 360
caggagatct accaggagct gacccggctg aaggctgcag tgggtgagct tccagagaaa 420
tctaagatgc aggagatcta ccaggagctg actcggctga aggctgcagt gggtgagctt 480
ccagagaaat ctaagatgca ggagatctac caggagctga ctccggctgaa ggctgcagtg 540
ggtgagcttc cagagaaatc taagcagcag gagatctacc aggagctgac ccagctgaag 600
gctgcagtgg gtgagcttcc agagaaatct aagcagcagg agatctacca ggagctgacc 660
cagctgaagg ctgcagtggg tgagcttcca gagaaatcta agcagcagga gatctaccag 720
gagctgaccc ggctgaaggc tgcaagtggaa cgccctgtgcc gccgctgccc ctgggaatgg 780
acattcttcc aaggaaactg ttacttcatg tctaactccc agcggaaactg gcacgactcc 840
atcaactgcct gcaaagaagt gggggcccaag ctgcgtcaa tcaaaagtgc tgaggagcag 900
aacttcctac agctgcagtc ttccagaagt aaccgcttca cctggatggg actttcagat 960
ctaaatgagg aaggcatgtg gcaatgggtg gacggctcac ctctgttgcc cagcttcaac 1020
cagtaytgga acagaggaga gcccaacaac gttggggagg aagactgcgc ggaatttagt 1080
ggcaatggct ggaatgacga caaatgtaat cttgc当地 tctggatctg caaaaagtcc 1140
gcagcctcct gctccaggaa tgaagaacag tttcttcctc cagccccctgc caccccaaac 1200
ccccctcctg cg 1212

<210> 11
<211> 1212
<212> DNA
<213> Gorilla gorilla

DRAFT genome sequence

<400> 11
atgagtgact ccaaggaacc aagactgcag cagctggcc tcctggagga ggaacagctg 60
agaggccttg gattccgaca gactcgaggc tacaagagct tagcagggtg tcttggccat 120
ggtccccctgg tgctgcaact cctctccttc acgctcttgg ctgcgctcct tgtccaagtg 180
tccaaggtcc ccagctccat aagtcaggaa caatccaggc aagacgcgat ctaccagaac 240
ctgacccagt ttaaagctgc agtgggttag ctctcagaga aatccaagct gcaggagatc 300
tatcaggagc tgacccagct gaaggctgca gtgggtgagc ttccagagaa atctaagcag 360
caggagatct accaggagct gagccagctg aaggctgcag tgggtgagct tccagagaaa 420
tctaaagcagc aggagatcta ccaggagctg acccggtctga aggctgcagt gggtgagctt 480
ccagagaaat ctaagcagca ggagatctac caggagctga cccggctgaa ggctgcagtg 540
ggtgagcttc cagagaaatc taagcagcag gagatctacc aggagcttag ccagctgaag 600
gctgcagtgg gtgagcttcc agagaaatct aagcagcagg agatctacca ggagctgagc 660
cagctgaagg ctgcagtggg tgagcttcca gagaatcta agcagcagga gatctaccag 720
gagctgaccc agctgaaggc tgcagtggaa cgccctgtgcc gccgctgccc ctggaatgg 780
acattcttcc aaggaaactg ttacttcatg tctaactccc agcggaaactg gcacgactcc 840
atcaccgcct gccaagaagt gggggcccaag ctcgtcgtaa tcaaaagtgc tgaggagcag 900
aacttcctac agctgcagtc ttccagaagt aaccgcttca cctggatgg actttcagat 960
ctaaatcatg aaggcacgtg gcaatgggtg gacggctcac ctctgttgcc cagttcgag 1020
cagtatttggaa acagaggaga gcccaacaac gttggggagg aagactgcgc ggaatttagt 1080
ggcaatggct ggaacgatga caaatgtaat ctgcacaaat tctggatctg caaaaagtct 1140
gcagcctcct gctccaggaa tgaagaacag tttcttcctc cagcctctgc cacccaaac 1200
ccccctcctg cg 1212

<210> 12
<211> 105
<212> PRT
<213> Pan troglodytes

<400> 12
Ser Leu Gln Cys Tyr Asn Cys Pro Asn Pro Thr Ala Asp Cys Lys Thr
1 5 10 15

Ala Val Asn Cys Ser Ser Asp Phe Asp Ala Cys Leu Ile Thr Lys Ala
20 25 30

Gly Leu Gln Val Tyr Asn Lys Cys Trp Lys Phe Glu His Cys Asn Phe
35 40 45

Asn Asp Val Thr Thr Arg Leu Arg Glu Asn Glu Leu Thr Tyr Tyr Cys
50 55 60

Cys Lys Lys Asp Leu Cys Asn Phe Asn Glu Gln Leu Glu Asn Gly Gly
65 70 75 80

Thr Ser Leu Ser Glu Lys Thr Val Leu Leu Leu Val Thr Pro Phe Leu
85 90 95

Ala Ala Ala Ala Trp Ser Leu His Pro
100 105

<210> 13
<211> 121
<212> PRT
<213> Pan troglodytes

<400> 13
Ser Leu Gln Cys Tyr Asn Cys Pro Asn Pro Thr Ala Asp Cys Lys Thr
1 5 10 15

Ala Val Asn Cys Ser Ser Asp Phe Asp Ala Cys Leu Ile Thr Lys Ala
20 25 30

Gly Leu Gln Val Tyr Asn Lys Cys Trp Lys Leu Glu His Cys Asn Phe
35 40 45

Lys Asp Leu Thr Thr Arg Leu Arg Glu Asn Glu Leu Thr Tyr Tyr Cys
50 55 60

Cys Lys Lys Asp Leu Cys Asn Phe Asn Glu Gln Leu Glu Asn Gly Gly
65 70 75 80

Asn Glu Gln Leu Glu Asn Gly Gly Asn Glu Gln Leu Glu Asn Gly Gly
85 90 95

Thr Ser Leu Ser Glu Lys Thr Val Leu Leu Arg Val Thr Pro Phe Leu
100 105 110

Ala Ala Ala Ala Trp Ser Leu His Pro
115 120

<210> 14
<211> 5140
<212> DNA
<213> Homo sapiens

<400> 14
ctccagacct acccagaaaag atgcccggat ggatcctgca gctccgtggc ttttctggga 60
agcagcggcc cctgctctca agagaccctg gtcctgtatg gtggccccaa gttgccagc 120
tggtgctagg gactcaggac agtttcccag aaaaggccaa gcgggcagcc ctcaggggg 180
ccgggtgagg aagctgggg gtcggaggc cacactgggt ccctgaaccc ctcgttgtt 240
tacagtgcag ctcccaagt ccacagacgt gggccggcac agcctcctgt acctgaagga 300
aatcgccgt ggctggttcg ggaagggttt cctggggag gtgaactctg gcatcagcag 360
tgcccagggtg gtggtaagg agctgcaggc tagtgcaggc gtgcaggagc agatgcagtt 420
cctggaggag gtgcagccct acagggccct gaagcacagc aacctgctcc agtgcctggc 480
ccagtgcgcc gaggtgacgc cctacctgct ggtgatggag ttctgcccac tggggaccc 540
caagggtac ctgcggagct gccgggtggc ggagtccatg gtcggacc cccggaccct 600
gcagcgcattt gctgtgagg tggcctgtgg cgtcctgcac cttcatcgca acaatttcgt 660
gcacagcgac ctggccctgc ggaactgcct gtcacggct gacgtacgg tgaagattgg 720
tgactatggc ctggctcaact gcaagtacag agaggactac ttcgtactg ccgaccagct 780
gtgggtgcct ctgcgtgga tcgcgcaga gctggggac gaggtgcata gcaacctgct 840
cgtcgtggac cagaccaaga gcggaaatgt gtggccctg ggcgtgacca tctggagct 900
ctttgagctg ggcacgcagc cctatccccca gcactcgac cagcagggtgc tggcgtacac 960
ggtcggggag cagcagctca agctgccccaa gccccagctg cagctgaccc tgcggaccg 1020
ctggtaacgag gtgtatgcagt tctgctggct gcagcccgag cagcggccca cagccgagga 1080

ggtgtcacctg ctgtgtcct acctgtgtgc caagggcgcc accgaagcag aggaggagtt 1140
tgaacggcgc tggcgctctc tgccggcccg cgaaaaaaggc gtggggccccg ggcccggtgc 1200
ggcgccccccc atgctggcg gcgtggtgga gctcgccgct gcctcgctct tccccgtgct 1260
ggagcagttc gcgggcgacg gcttccacgc ggacggcgac gacgtgctga cggtgaccga 1320
gaccagccga ggcctaatt ttgagtacaa gtgggaggcg ggccgcggcg cggaggcctt 1380
cccgccacg ctgagccctg gccgcaccgc acgcctgcag gagctgtgcg ccccccacgg 1440
cgcgcccccg ggcgtggttc cggtgctcag cgccacacgc ccgtcgctgg gcagcgagta 1500
cttcatccgc ctagaggagg ccgcacccgc cgccggccac gaccctgact ggcggcgtg 1560
cgcccccaat ccacctgcca ccgcggacca ggacgcacgc tctgacggca gcaccgcgc 1620
ctcgctggcc atggagccgc tgctggcca cggccaccc gtcgacgtcc cctggggccg 1680
cgccgaccac taccctcgca gaagcttggc gcgggacccg ctctgcccct cacgctctcc 1740
ctcgccctcg gcggggccccc tgagtctggc ggagggagga goggaggatg cagactgggg 1800
cgtggccgccc ttctgtcctg ctttcttcga ggaccactg ggcacgtccc ctttggggag 1860
ctcagggcgcc ccccccgtgc cgctgactgg cgaggatgag ctagaggagg tgggagcgcg 1920
gagggccgccc cagcgccggc actggcgctc caacgtgtca gccaacaaca acagcgccag 1980
ccgctgtcca gagtcctggg accccgtctc tgccggctgc cacgctgagg gctgccccag 2040
tccaaagcag accccacggg cctccccgaa gccggggtaa cctggagagc ctctgcttgg 2100
gctccaggca gcctctgccc aggagccagg ctgctgcccc ggcctccctc atctatgctc 2160
tgcccaggc ctggcacctg ctccctgcct gttacaccc tcctggacag agacagccag 2220
tagtgggggt gaccacccgc aggcagagcc caagcttgcc acggaggctg agggcactac 2280
cgaccccgcc ctgcccccttc ctcccgatcc ctcccatcc caggaggag ccccaattcc 2340
ctcggaggag gccagtgccc ccgacgcccc tggatgcctg cctgactctc ccacgcctgc 2400
tactggtggc gaggtgtctg ccatcaagct ggcttctgcc ctgaatggca gcagcagctc 2460
tccccgggtg gaggcaccca gcagtggagga tgaggacacg gctgaggccca cctcaggcat 2520

TOEGLER ET AL.

cttcaccgac acgtccagcg acggcctgca ggccaggagg ccggatgtgg tgccagcctt 2580
ccgctctctg cagaagcagg tggggacccc cgactccctg gactccctgg acatcccgac 2640
ctcagccagt gatggtggt atgaggtctt cagcccgtcg gccactggcc cctctggagg 2700
gcagccgcga ggcgtggaca gtggctatga caccgagaac tatgagtccc ctgagttgt 2760
gctcaaggag gcgcaggaag ggtgtgagcc ccaggcctt gcggagctgg cctcagaggg 2820
tgagggcccc gggcccgaga cacggctctc caccccttc agtggcctca acgagaagaa 2880
tccctaccga gactctgcct acttctcaga cctcgaggct gaggccgagg ccacccagg 2940
cccagagaag aagtgcggcg gggaccgagc ccccgccca gagctgggcc tgccgagcac 3000
tggcagccg tctgagcagg tctgtctcag gcctgggtt tccggggagg cacaaggctc 3060
tggccccggg gaggtgtgc ccccactgct gcagcttgaa gggtcctccc cagagccag 3120
cacctgcccc tcgggcctgg tcccagagcc tccggagccc caaggcccag ccaaggtgcg 3180
gcctggggcc agccccagct gctccagtt tttcctgctg accccgggtc cgctgagatc 3240
agaaggcaac agctctgagt tccaggggcc cccaggactg ttgtcagggc cggcccccaca 3300
aaagcggatg gggggcccaag gcaccccaag agccccactc cgctggctc tgccggcct 3360
ccctgcggcc ttggagggcc ggccggagga ggaggaggag gacagtgagg acagcgacga 3420
gtctgacgag gagctccgct gctacagcggt ccaggagcct agcgaggaca gcgaagagga 3480
ggcgccggcg gtgcccgtgg tggtggtga gagccagagc gcgcgcaacc tgcgcagcct 3540
gctcaagatg cccagcctgc tgtccgagac cttctgcgag gacctggaaac gcaagaagaa 3600
ggccgtgtcc ttcttcgacg acgtcaccgt ctacctctt gaccagaaaa gccccacccg 3660
ggagctcggg gagcccttcc cggcgccaa ggaatcgccc cctacgttcc ttagggggag 3720
ccccggctct cccagcgccc ccaaccggcc gcagcaggct gatggctccc caaatggctc 3780
cacagcggaa gagggtgttg ggttcgcgtg ggacgacgac ttcccgtga tgacggccaa 3840
ggcagccttc gccatggccc tagacccggc cgacccggcc cggctgcgc ccacgcccac 3900
ccccgctccc ttctcgcgct tcacggtgac gcccggccc achtcccgt tctccatcac 3960

gcacgtgtct gactcggacg ccgagtccaa gagaggacct gaagctggtg ccgggggtga 4020
gagtaaagag gctttagagacc tgggcagctc ctgcccctca aggctggcgt caccggagcc 4080
cctgccaggc agcagcgagg atggtgaccg agaaggtggg gaccacgtcc tggtggtgt 4140
tggcagcaga ttccaggtgcc tctgccccac gcgggtgtct ggagaagccc gtgggatgag 4200
aggccctgga tggtagatcg gccatgctcc gccccagagg cagaattcgt ctgggcttt 4260
aggcttgctg ctagccctg ggggcgcctg gagccacagt gggtgtctgt acacacatac 4320
acactcaaaa ggggcccagtg cccctggca cggcggcccc caccctctgc cctgcctgcc 4380
tggcctcgga ggacccgcat gccccatccg gcagctcctc cggtgtgctc acaggacact 4440
taaaccagga cgaggcatgg ccccgagaca ctggcagggtt tgtgagcctc ttcccacccc 4500
ctgtgcccccc acccttgcct gttcctggt ggctcagggc aaggagtggc cttggcgcc 4560
cgtgtcggtc ctgtttccgc tgcccttatac tcaaagtccg tggctgtttc cccttcactg 4620
actcagctag acccgtaagc ccacccttcc cacagggAAC aggctgctcc cacctgggtc 4680
ccgctgtggc cacgggtggc agccaaaaag atcaggggtg gaggggcttc caggctgtac 4740
tcctgccccg tgggccccgt tctagaggtg ccattggcag gaccgtgcag gcagctcccc 4800
tctgtggggc agtatctggt cctgtgcccc agctgccaaa ggagagtggg ggccatgccc 4860
cgcagtcagt gttggggggc tcctgcctac agggagaggg atggtggggg aggggtggag 4920
ctggggcag ggcagcacag ggaatattt tgtaactaac taactgctgt gttggagcg 4980
aatggaaagtt gggtgatttt aagttattgt tgccaaagag atgtaaagtt tattgttgct 5040
tcgcaggggg atttgtttt tgttttgtt gaggcttaga acgctggtgc aatgtttct 5100
tgttccttgtt ttttaagag aaatgaagct aagaaaaaag 5140

<210> 15
<211> 5140
<212> DNA
<213> Homo sapiens

<220>

<221> CDS
 <222> (413)..(4036)

<400> 15
 ctccagacct acccagaaaag atgcccggat ggatcctgca gctccgtggc ttttctggga 60
 agcagcggcc cctgctctca agagaccctg gtcctgatg gtggccccaa gttgccagc 120
 tggtgctagg gactcaggac agtttcccag aaaaggccaa gcgggcagcc ctcaggggg 180
 ccgggtgagg aagctggggg gtgcggaggc cacactgggt ccctgaaccc ctcgttgg 240
 tacagtgcag ctcctaagt ccacagacgt gggccggcac agcctcctgt acctgaagga 300
 aatcgccgt ggctggttcg ggaagggttt cctggggag gtgaactctg gcatcagcag 360
 tgcccaggtg gtggtaagg agctgcaggc tagtgcacgc gtgcaggagc ag atg cag 418
 Met Gln
 1

ttc ctg gag gag gtg cag ccc tac agg gcc ctg aag cac agc aac ctg 466
 Phe Leu Glu Glu Val Gln Pro Tyr Arg Ala Leu Lys His Ser Asn Leu
 5 10 15

ctc cag tgc ctg gcc cag tgc gcc gag gtg acg ccc tac ctg ctg gtg 514
 Leu Gln Cys Leu Ala Gln Cys Ala Glu Val Thr Pro Tyr Leu Leu Val
 20 25 30

atg gag ttc tgc cca ctg ggg gac ctc aag ggc tac ctg cgg agc tgc 562
 Met Glu Phe Cys Pro Leu Gly Asp Leu Lys Gly Tyr Leu Arg Ser Cys
 35 40 45 50

cggtg gcg gag tcc atg gct ccc gac ccc cgg acc ctg cag cgc atg 610
 Arg Val Ala Glu Ser Met Ala Pro Asp Pro Arg Thr Leu Gln Arg Met
 55 60 65

gcc tgt gag gtg gcc tgt ggc gtc ctg cac ctt cat cgc aac aat ttc 658
 Ala Cys Glu Val Ala Cys Gly Val Leu His Leu His Arg Asn Asn Phe
 70 75 80

gtg cac agc gac ctg gcc ctg cgg aac tgc ctg ctc acg gct gac ctg 706
 Val His Ser Asp Leu Ala Leu Arg Asn Cys Leu Leu Thr Ala Asp Leu
 85 90 95

acg gtg aag att ggt gac tat ggc ctg gct cac tgc aag tac aga gag 754
 Thr Val Lys Ile Gly Asp Tyr Gly Leu Ala His Cys Lys Tyr Arg Glu
 100 105 110

DRAFT

gac tac ttc gtg act gcc gac cag ctg tgg gtg cct ctg cgc tgg atc		802
Asp Tyr Phe Val Thr Ala Asp Gln Leu Trp Val Pro Leu Arg Trp Ile		
115	120	125
130		
gcg cca gag ctg gtg gac gag gtg cat aac aac ctg ctc gtc gtg gac		850
Ala Pro Glu Leu Val Asp Glu Val His Ser Asn Leu Leu Val Val Asp		
135	140	145
cag acc aag agc ggg aat gtg tgg tcc ctg ggc gtg acc atc tgg gag		898
Gln Thr Lys Ser Gly Asn Val Trp Ser Leu Gly Val Thr Ile Trp Glu		
150	155	160
ctc ttt gag ctg ggc acg cag ccc tat ccc cag cac tcg gac cag cag		946
Leu Phe Glu Leu Gly Thr Gln Pro Tyr Pro Gln His Ser Asp Gln Gln		
165	170	175
gtg ctg gcg tac acg gtc cgg gag cag cag ctc aag ctg ccc aag ccc		994
Val Leu Ala Tyr Thr Val Arg Glu Gln Gln Leu Lys Leu Pro Lys Pro		
180	185	190
cag ctg cag ctg acc ctg tcg gac cgc tgg tac gag gtg atg cag ttc		1042
Gln Leu Gln Leu Thr Leu Ser Asp Arg Trp Tyr Glu Val Met Gln Phe		
195	200	205
210		
tgc tgg ctg cag ccc gag cag cgg ccc aca gcc gag gag gtg cac ctg		1090
Cys Trp Leu Gln Pro Glu Gln Arg Pro Thr Ala Glu Glu Val His Leu		
215	220	225
ctg ctg tcc tac ctg tgt gcc aag ggc gcc acc gaa gca gag gag gag		1138
Leu Leu Ser Tyr Leu Cys Ala Lys Gly Ala Thr Glu Ala Glu Glu Glu		
230	235	240
ttt gaa cgg cgc tgg cgc tct ctg cgg ccc ggc ggg ggc ggc gtg ggg		1186
Phe Glu Arg Arg Trp Arg Ser Leu Arg Pro Gly Gly Gly Val Gly		
245	250	255
ccc ggg ccc ggt gcg gcg ggg ccc atg ctg ggc ggc gtg gtg gag ctc		1234
Pro Gly Pro Gly Ala Ala Gly Pro Met Leu Gly Gly Val Val Glu Leu		
260	265	270
gcc gct gcc tcg tcc ttc ccg ctg ctg gag cag ttc gcg ggc gac ggc		1282
Ala Ala Ala Ser Ser Phe Pro Leu Leu Glu Gln Phe Ala Gly Asp Gly		
275	280	285
290		
ttc cac gcg gac gac gac gtg ctg acg gtg acc gag acc agc cga		1330
Phe His Ala Asp Gly Asp Asp Val Leu Thr Val Thr Glu Thr Ser Arg		
295	300	305

TIGEESDSESSS

ggc ctc aat ttt gag tac aag tgg gag gcg ggc cgc ggc gag gcc Gly Leu Asn Phe Glu Tyr Lys Trp Glu Ala Gly Arg Gly Ala Glu Ala 310 315 320	1378
ttc ccg gcc acg ctg agc cct ggc cgc acc gca cgc ctg cag gag ctg Phe Pro Ala Thr Leu Ser Pro Gly Arg Thr Ala Arg Leu Gln Glu Leu 325 330 335	1426
tgc gcc ccc gac ggc gcg ccc ccg ggc gtg gtt ccg gtg ctc agc gcg Cys Ala Pro Asp Gly Ala Pro Pro Gly Val Val Pro Val Leu Ser Ala 340 345 350	1474
cac agc ccg tcg ctg ggc agc gag tac ttc atc cgc cta gag gag gcc His Ser Pro Ser Leu Gly Ser Glu Tyr Phe Ile Arg Leu Glu Glu Ala 355 360 365 370	1522
gca ccc gcc ggc cac gac cct gac tgc gcc ggc tgc gcc ccc agt Ala Pro Ala Ala Gly His Asp Pro Asp Cys Ala Gly Cys Ala Pro Ser 375 380 385	1570
cca cct gcc acc gcg gac cag gac gac tct gac ggc agc acc gcc Pro Pro Ala Thr Ala Asp Gln Asp Asp Ser Asp Gly Ser Thr Ala 390 395 400	1618
gcc tcg ctg gcc atg gag ccg ctg ggc cac ggg cca ccc gtc gac Ala Ser Leu Ala Met Glu Pro Leu Leu Gly His Gly Pro Pro Val Asp 405 410 415	1666
gtc ccc tgg ggc cgc ggc gac cac tac cct cgc aga agc ttg gcg cgg Val Pro Trp Gly Arg Gly Asp His Tyr Pro Arg Arg Ser Leu Ala Arg 420 425 430	1714
gac ccg ctc tgc ccc tca cgc tct ccc tcg ccc tcg gcg ggg ccc ctg Asp Pro Leu Cys Pro Ser Arg Ser Pro Ser Pro Ser Ala Gly Pro Leu 435 440 445 450	1762
agt ctg gcg gag gga gga gcg gag gat gca gac tgg ggc gtg gcc gcc Ser Leu Ala Glu Gly Gly Ala Glu Asp Ala Asp Trp Gly Val Ala Ala 455 460 465	1810
ttc tgt cct gcc ttc ttc gag gac cca ctg ggc acg tcc cct ttg ggg Phe Cys Pro Ala Phe Phe Glu Asp Pro Leu Gly Thr Ser Pro Leu Gly 470 475 480	1858
agc tca ggg gcg ccc ccg ctg ccg act ggc gag gat gag cta gag Ser Ser Gly Ala Pro Pro Leu Pro Leu Thr Gly Glu Asp Glu Leu Glu 485 490 495	1906

gag gtg gga gcg cgg agg gcc gcc cag cgc ggg cac tgg cgc tcc aac			1954
Glu Val Gly Ala Arg Arg Ala Ala Gln Arg Gly His Trp Arg Ser Asn			
500	505	510	
 gtg tca gcc aac aac aac agc ggc agc cgc tgt cca gag tcc tgg gac			2002
Val Ser Ala Asn Asn Asn Ser Gly Ser Arg Cys Pro Glu Ser Trp Asp			
515	520	525	530
 ccc gtc tct gcg ggc tgc cac gct gag ggc tgc ccc agt cca aag cag			2050
Pro Val Ser Ala Gly Cys His Ala Glu Gly Cys Pro Ser Pro Lys Gln			
535	540	545	
 acc cca cgg gcc tcc ccc gag ccg ggg tac cct gga gag cct ctg ctt			2098
Thr Pro Arg Ala Ser Pro Glu Pro Gly Tyr Pro Gly Glu Pro Leu Leu			
550	555	560	
 ggg ctc cag gca gcc tct gcc cag gag cca ggc tgc tgc ccc ggc ctc			2146
Gly Leu Gln Ala Ala Ser Ala Gln Glu Pro Gly Cys Cys Pro Gly Leu			
565	570	575	
 cct cat cta tgc tct gcc cag ggc ctg gca cct gct ccc tgc ctg gtt			2194
Pro His Leu Cys Ser Ala Gln Gly Leu Ala Pro Ala Pro Cys Leu Val			
580	585	590	
 aca ccc tcc tgg aca gag aca gcc agt agt ggg ggt gac cac ccg cag			2242
Thr Pro Ser Trp Thr Glu Thr Ala Ser Ser Gly Gly Asp His Pro Gln			
595	600	605	610
 gca gag ccc aag ctt gcc acg gag gct gag ggc act acc gga ccc cgc			2290
Ala Glu Pro Lys Leu Ala Thr Glu Ala Glu Gly Thr Thr Gly Pro Arg			
615	620	625	
 ctg ccc ctt cct tcc gtc ccc tcc cca tcc cag gag gga gcc cca ctt			2338
Leu Pro Leu Pro Ser Val Pro Ser Pro Ser Gln Glu Gly Ala Pro Leu			
630	635	640	
 ccc tcg gag gag gcc agt gcc ccc gac gcc cct gat gcc ctg cct gac			2386
Pro Ser Glu Glu Ala Ser Ala Pro Asp Ala Pro Asp Ala Leu Pro Asp			
645	650	655	
 tct ccc acg cct gct act ggt ggc gag gtg tct gcc atc aag ctg gct			2434
Ser Pro Thr Pro Ala Thr Gly Gly Glu Val Ser Ala Ile Lys Leu Ala			
660	665	670	
 tct gcc ctg aat ggc agc agc agc tct ccc gag gtg gag gca ccc agc			2482
Ser Ala Leu Asn Gly Ser Ser Ser Pro Glu Val Glu Ala Pro Ser			
675	680	685	690

agt gag gat gag gac acg gct gag gcc acc tca ggc atc ttc acc gac			2530
Ser Glu Asp Glu Asp Thr Ala Glu Ala Thr Ser Gly Ile Phe Thr Asp			
695	700	705	
acg tcc agc gac ggc ctg cag gcc agg agg ccg gat gtg gtg cca gcc			2578
Thr Ser Ser Asp Gly Leu Gln Ala Arg Arg Pro Asp Val Val Pro Ala			
710	715	720	
ttc cgc tct ctg cag aag cag gtg ggg acc ccc gac tcc ctg gac tcc			2626
Phe Arg Ser Leu Gln Lys Gln Val Gly Thr Pro Asp Ser Leu Asp Ser			
725	730	735	
ctg gac atc ccg tcc tca gcc agt gat ggt ggc tat gag gtc ttc agc			2674
Leu Asp Ile Pro Ser Ser Ala Ser Asp Gly Gly Tyr Glu Val Phe Ser			
740	745	750	
ccg tcg gcc act ggc ccc tct gga ggg cag ccg cga gcg ctg gac agt			2722
Pro Ser Ala Thr Gly Pro Ser Gly Gly Gln Pro Arg Ala Leu Asp Ser			
755	760	765	770
ggc tat gac acc gag aac tat gag tcc cct gag ttt gtg ctc aag gag			2770
Gly Tyr Asp Thr Glu Asn Tyr Glu Ser Pro Glu Phe Val Leu Lys Glu			
775	780	785	
gcg cag gaa ggg tgt gag ccc cag gcc ttt gcg gag ctg gcc tca gag			2818
Ala Gln Glu Gly Cys Glu Pro Gln Ala Phe Ala Glu Leu Ala Ser Glu			
790	795	800	
ggt gag ggc ccc ggg ccc gag aca cgg ctc tcc acc tcc ctc agt ggc			2866
Gly Glu Gly Pro Gly Pro Glu Thr Arg Leu Ser Thr Ser Leu Ser Gly			
805	810	815	
ctc aac gag aag aat ccc tac cga gac tct gcc tac ttc tca gac ctc			2914
Leu Asn Glu Lys Asn Pro Tyr Arg Asp Ser Ala Tyr Phe Ser Asp Leu			
820	825	830	
gag gct gag gcc gag gcc acc tca ggc cca gag aag aag tgc ggc ggg			2962
Glu Ala Glu Ala Glu Ala Thr Ser Gly Pro Glu Lys Lys Cys Gly Gly			
835	840	845	850
gac cga gcc ccc ggg cca gag ctg ggc ctg ccg agc act ggg cag ccg			3010
Asp Arg Ala Pro Gly Pro Glu Leu Gly Leu Pro Ser Thr Gly Gln Pro			
855	860	865	
tct gag cag gtc tgt ctc agg cct ggg gtt tcc ggg gag gca caa ggc			3058
Ser Glu Gln Val Cys Leu Arg Pro Gly Val Ser Gly Glu Ala Gln Gly			
870	875	880	

tct ggc ccc ggg gag gtg ctg ccc cca ctg ctg cag ctt gaa ggg tcc		3106
Ser Gly Pro Gly Glu Val Leu Pro Pro Leu Leu Gln Leu Glu Gly Ser		
885	890	895
tcc cca gag ccc agc acc tgc ccc tcg ggc ctg gtc cca gag cct ccg		3154
Ser Pro Glu Pro Ser Thr Cys Pro Ser Gly Leu Val Pro Glu Pro Pro		
900	905	910
gag ccc caa ggc cca gcc aag gtg cgg cct ggg ccc agc ccc agc tgc		3202
Glu Pro Gln Gly Pro Ala Lys Val Arg Pro Gly Pro Ser Pro Ser Cys		
915	920	925
930		
tcc cag ttt ttc ctg ctg acc ccg gtt ccg ctg aga tca gaa ggc aac		3250
Ser Gln Phe Phe Leu Leu Thr Pro Val Pro Leu Arg Ser Glu Gly Asn		
935	940	945
agc tct gag ttc cag ggg ccc cca gga ctg ttg tca ggg ccg gcc cca		3298
Ser Ser Glu Phe Gln Gly Pro Pro Gly Leu Leu Ser Gly Pro Ala Pro		
950	955	960
caa aag cgg atg ggg ggc cca ggc acc ccc aga gcc cca ctc cgc ctg		3346
Gln Lys Arg Met Gly Gly Pro Gly Thr Pro Arg Ala Pro Leu Arg Leu		
965	970	975
gct ctg ccc ggc ctc cct gcg gcc ttg gag ggc cgg ccg gag gag gag		3394
Ala Leu Pro Gly Leu Pro Ala Ala Leu Glu Gly Arg Pro Glu Glu Glu		
980	985	990
gag gag gac agt gag gac agc gac gag tct gac gag gag ctc cgc tgc		3442
Glu Glu Asp Ser Glu Asp Ser Asp Glu Ser Asp Glu Glu Leu Arg Cys		
995	1000	1005
1010		
tac agc gtc cag gag cct agc gag gac agc gaa gag gag gcg ccg gcg		3490
Tyr Ser Val Gln Glu Pro Ser Glu Asp Ser Glu Glu Ala Pro Ala		
1015	1020	1025
gtg ccc gtg gtg gct gag agc cag agc gcg cgc aac ctg cgc agc		3538
Val Pro Val Val Val Ala Glu Ser Gln Ser Ala Arg Asn Leu Arg Ser		
1030	1035	1040
ctg ctc aag atg ccc agc ctg ctg tcc gag acc ttc tgc gag gac ctg		3586
Leu Leu Lys Met Pro Ser Leu Leu Ser Glu Thr Phe Cys Glu Asp Leu		
1045	1050	1055
gaa cgc aag aag aag gcc gtg tcc ttc ttc gac gac gtc acc gtc tac		3634
Glu Arg Lys Lys Lys Ala Val Ser Phe Phe Asp Asp Val Thr Val Tyr		
1060	1065	1070

卷之三

ctc ttt gac cag gaa agc ccc acc cgg gag ctc ggg gag ccc ttc ccg 3682
 Leu Phe Asp Gln Glu Ser Pro Thr Arg Glu Leu Gly Glu Pro Phe Pro
 1075 1080 1085 1090

 ggc gcc aag gaa tcg ccc cct acg ttc ctt agg ggg agc ccc ggc tct 3730
 Gly Ala Lys Glu Ser Pro Pro Thr Phe Leu Arg Gly Ser Pro Gly Ser
 1095 1100 1105

 ccc agc gcc ccc aac cgg ccg cag cag gct gat ggc tcc cca aat ggc 3778
 Pro Ser Ala Pro Asn Arg Pro Gln Gln Ala Asp Gly Ser Pro Asn Gly
 1110 1115 1120

 tcc aca gcg gaa gag ggt ggt ggg ttc gcg tgg gac gac gac ttc ccg 3826
 Ser Thr Ala Glu Glu Gly Gly Phe Ala Trp Asp Asp Asp Phe Pro
 1125 1130 1135

 ctg atg acg gcc aag gca gcc ttc gcc atg gcc cta gac ccg gcc gca 3874
 Leu Met Thr Ala Lys Ala Ala Phe Ala Met Ala Leu Asp Pro Ala Ala
 1140 1145 1150

 ccc gcc ccg gct gcg ccc acg ccc acg ccc gct ccc ttc tcg ccg ttc 3922
 Pro Ala Pro Ala Ala Pro Thr Pro Thr Pro Ala Pro Phe Ser Arg Phe
 1155 1160 1165 1170

 acg gtg tcg ccc gcg ccc acg tcc cgc ttc tcc atc acg cac gtg tct 3970
 Thr Val Ser Pro Ala Pro Thr Ser Arg Phe Ser Ile Thr His Val Ser
 1175 1180 1185

 gac tcg gac gcc gag tcc aag aga gga cct gaa gct ggt gcc ggg ggt 4018
 Asp Ser Asp Ala Glu Ser Lys Arg Gly Pro Glu Ala Gly Ala Gly Gly
 1190 1195 1200

 gag agt aaa gag gct tga gacccctggca gtcctgtccc ctcaaggctg 4066
 Glu Ser Lys Glu Ala
 1205

 gcgtcacccgg agccccctgcc aggcagcagc gaggatggtg accgagaagg tggggaccac 4126

 gtcctggtgg ctgttggcag cagattcagg tgcctctgcc ccacgcggtg tcctggagaa 4186

 gcccgtggga tgagaggccc tggatggtag atcggccatg ctccgccccca gaggcagaat 4246

 tcgtctgggc tttaggctt gctgctagcc cctggggccg cctggagcca cagtgggtgt 4306

 ctgtacacac atacacactc aaaaggggcc agtgccccctg ggcacggcgg cccccaccct 4366

 ctqccctqcc tqccctqqcc cqqaqqaccc qcatqcccca tccqqcaqct cctccqqtgt 4426

gctcacagga cacttaaacc aggacgaggg atggcccgaa gacactggca ggtttgttag 4486
cctcttccca cccccctgtgc ccccaccctt gcctggttcc tggggctca gggcaaggag 4546
tgccccctggg cgccccgtgtc ggtcctgttt ccgctgccct tatctcaaag tccgtggctg 4606
tttcccccttc actgactcag ctagaccgt aagcccaccc ttcccacagg gaacaggctg 4666
ctccccacctg ggtccccgtg tggccacggg gggcagccca aaagatcagg ggtggagggg 4726
cttccaggct gtactcctgc cccgtggggcc ccgttctaga ggtgcccttg gcaggaccgt 4786
gcaggcagct cccctctgtg gggcagtatac tggtcctgtg ccccagctgc caaaggagag 4846
tggggccat gccccgcagt cagtgttggg gggctcctgc ctacagggag agggatggtg 4906
gggaaggggt ggagctgggg gcagggcagc acagggaaata tttttgttaac taactaactg 4966
ctgtggttgg agcgaatggaa agttgggtga ttttaagtta ttgttgc当地 agagatgtaa 5026
agtttattgt tgcttcgcag ggggatttgt tttgtgtttt gttttagggct tagaacgctg 5086
gtgcaatgtt ttcttggcc ttgtttttta agagaaatga agctaagaaa aaag 5140

<210> 16
<211> 1207
<212> PRT
<213> Homo sapiens

<400> 16
Met Gln Phe Leu Glu Glu Val Gln Pro Tyr Arg Ala Leu Lys His Ser
1 5 10 15

Asn Leu Leu Gln Cys Leu Ala Gln Cys Ala Glu Val Thr Pro Tyr Leu
20 25 30

Leu Val Met Glu Phe Cys Pro Leu Gly Asp Leu Lys Gly Tyr Leu Arg
35 40 45

Ser Cys Arg Val Ala Glu Ser Met Ala Pro Asp Pro Arg Thr Leu Gln
50 55 60

Arg Met Ala Cys Glu Val Ala Cys Gly Val Leu His Leu His Arg Asn
65 70 75 80

Asn Phe Val His Ser Asp Leu Ala Leu Arg Asn Cys Leu Leu Thr Ala
85 90 95

Asp Leu Thr Val Lys Ile Gly Asp Tyr Gly Leu Ala His Cys Lys Tyr
100 105 110

Arg Glu Asp Tyr Phe Val Thr Ala Asp Gln Leu Trp Val Pro Leu Arg
115 120 125

Trp Ile Ala Pro Glu Leu Val Asp Glu Val His Ser Asn Leu Leu Val
130 135 140

Val Asp Gln Thr Lys Ser Gly Asn Val Trp Ser Leu Gly Val Thr Ile
145 150 155 160

Trp Glu Leu Phe Glu Leu Gly Thr Gln Pro Tyr Pro Gln His Ser Asp
165 170 175

Gln Gln Val Leu Ala Tyr Thr Val Arg Glu Gln Gln Leu Lys Leu Pro
180 185 190

Lys Pro Gln Leu Gln Leu Thr Leu Ser Asp Arg Trp Tyr Glu Val Met
195 200 205

Gln Phe Cys Trp Leu Gln Pro Glu Gln Arg Pro Thr Ala Glu Glu Val
210 215 220

His Leu Leu Leu Ser Tyr Leu Cys Ala Lys Gly Ala Thr Glu Ala Glu
225 230 235 240

Glu Glu Phe Glu Arg Arg Trp Arg Ser Leu Arg Pro Gly Gly Gly
245 250 255

Val Gly Pro Gly Pro Gly Ala Ala Gly Pro Met Leu Gly Gly Val Val
260 265 270

Glu Leu Ala Ala Ala Ser Ser Phe Pro Leu Leu Glu Gln Phe Ala Gly
275 280 285

Asp Gly Phe His Ala Asp Gly Asp Asp Val Leu Thr Val Thr Glu Thr
290 295 300

Ser Arg Gly Leu Asn Phe Glu Tyr Lys Trp Glu Ala Gly Arg Gly Ala
305 310 315 320

Glu Ala Phe Pro Ala Thr Leu Ser Pro Gly Arg Thr Ala Arg Leu Gln
325 330 335

Glu Leu Cys Ala Pro Asp Gly Ala Pro Pro Gly Val Val Pro Val Leu
340 345 350

Top sequence

Ser Ala His Ser Pro Ser Leu Gly Ser Glu Tyr Phe Ile Arg Leu Glu
355 360 365

Glu Ala Ala Pro Ala Ala Gly His Asp Pro Asp Cys Ala Gly Cys Ala
370 375 380

Pro Ser Pro Pro Ala Thr Ala Asp Gln Asp Asp Asp Ser Asp Gly Ser
385 390 395 400

Thr Ala Ala Ser Leu Ala Met Glu Pro Leu Leu Gly His Gly Pro Pro
405 410 415

Val Asp Val Pro Trp Gly Arg Gly Asp His Tyr Pro Arg Arg Ser Leu
420 425 430

Ala Arg Asp Pro Leu Cys Pro Ser Arg Ser Pro Ser Pro Ala Gly
435 440 445

Pro Leu Ser Leu Ala Glu Gly Gly Ala Glu Asp Ala Asp Trp Gly Val
450 455 460

Ala Ala Phe Cys Pro Ala Phe Phe Glu Asp Pro Leu Gly Thr Ser Pro
465 470 475 480

Leu Gly Ser Ser Gly Ala Pro Pro Leu Pro Leu Thr Gly Glu Asp Glu
485 490 495

Leu Glu Glu Val Gly Ala Arg Arg Ala Ala Gln Arg Gly His Trp Arg
500 505 510

Ser Asn Val Ser Ala Asn Asn Ser Gly Ser Arg Cys Pro Glu Ser
515 520 525

Trp Asp Pro Val Ser Ala Gly Cys His Ala Glu Gly Cys Pro Ser Pro
530 535 540

Lys Gln Thr Pro Arg Ala Ser Pro Glu Pro Gly Tyr Pro Gly Glu Pro
545 550 555 560

Leu Leu Gly Leu Gln Ala Ala Ser Ala Gln Glu Pro Gly Cys Cys Pro
565 570 575

Gly Leu Pro His Leu Cys Ser Ala Gln Gly Leu Ala Pro Ala Pro Cys
580 585 590

Leu Val Thr Pro Ser Trp Thr Glu Thr Ala Ser Ser Gly Gly Asp His
595 600 605

Pro Gln Ala Glu Pro Lys Leu Ala Thr Glu Ala Glu Gly Thr Thr Gly
610 615 620

Pro Arg Leu Pro Leu Pro Ser Val Pro Ser Pro Ser Gln Glu Gly Ala
625 630 635 640

Pro Leu Pro Ser Glu Glu Ala Ser Ala Pro Asp Ala Pro Asp Ala Leu
645 650 655

Pro Asp Ser Pro Thr Pro Ala Thr Gly Gly Glu Val Ser Ala Ile Lys
660 665 670

Leu Ala Ser Ala Leu Asn Gly Ser Ser Ser Pro Glu Val Glu Ala
675 680 685

Pro Ser Ser Glu Asp Glu Asp Thr Ala Glu Ala Thr Ser Gly Ile Phe
690 695 700

Thr Asp Thr Ser Ser Asp Gly Leu Gln Ala Arg Arg Pro Asp Val Val
705 710 715 720

Pro Ala Phe Arg Ser Leu Gln Lys Gln Val Gly Thr Pro Asp Ser Leu
725 730 735

Asp Ser Leu Asp Ile Pro Ser Ser Ala Ser Asp Gly Gly Tyr Glu Val
740 745 750

Phe Ser Pro Ser Ala Thr Gly Pro Ser Gly Gly Gln Pro Arg Ala Leu
755 760 765

Asp Ser Gly Tyr Asp Thr Glu Asn Tyr Glu Ser Pro Glu Phe Val Leu
770 775 780

Lys Glu Ala Gln Glu Gly Cys Glu Pro Gln Ala Phe Ala Glu Leu Ala
785 790 795 800

Ser Glu Gly Glu Gly Pro Gly Pro Glu Thr Arg Leu Ser Thr Ser Leu
805 810 815

Ser Gly Leu Asn Glu Lys Asn Pro Tyr Arg Asp Ser Ala Tyr Phe Ser
820 825 830

Asp Leu Glu Ala Glu Ala Glu Ala Thr Ser Gly Pro Glu Lys Lys Cys
835 840 845

Gly Gly Asp Arg Ala Pro Gly Pro Glu Leu Gly Leu Pro Ser Thr Gly
850 855 860

Gln Pro Ser Glu Gln Val Cys Leu Arg Pro Gly Val Ser Gly Glu Ala
865 870 875 880

Gln Gly Ser Gly Pro Gly Glu Val Leu Pro Pro Leu Leu Gln Leu Glu
885 890 895

Gly Ser Ser Pro Glu Pro Ser Thr Cys Pro Ser Gly Leu Val Pro Glu
900 905 910

Pro Pro Glu Pro Gln Gly Pro Ala Lys Val Arg Pro Gly Pro Ser Pro
915 920 925

Ser Cys Ser Gln Phe Phe Leu Leu Thr Pro Val Pro Leu Arg Ser Glu
930 935 940

Gly Asn Ser Ser Glu Phe Gln Gly Pro Pro Gly Leu Leu Ser Gly Pro
945 950 955 960

Ala Pro Gln Lys Arg Met Gly Gly Pro Gly Thr Pro Arg Ala Pro Leu
965 970 975

Arg Leu Ala Leu Pro Gly Leu Pro Ala Ala Leu Glu Gly Arg Pro Glu
980 985 990

Glu Glu Glu Asp Ser Glu Asp Ser Asp Glu Ser Asp Glu Glu Leu
995 1000 1005

Arg Cys Tyr Ser Val Gln Glu Pro Ser Glu Asp Ser Glu Glu Ala
1010 1015 1020

Pro Ala Val Pro Val Val Ala Glu Ser Gln Ser Ala Arg Asn Leu
1025 1030 1035 1040

Arg Ser Leu Leu Lys Met Pro Ser Leu Leu Ser Glu Thr Phe Cys Glu
1045 1050 1055

Asp Leu Glu Arg Lys Lys Ala Val Ser Phe Phe Asp Asp Val Thr
1060 1065 1070

Val Tyr Leu Phe Asp Gln Glu Ser Pro Thr Arg Glu Leu Gly Glu Pro
1075 1080 1085

Phe Pro Gly Ala Lys Glu Ser Pro Pro Thr Phe Leu Arg Gly Ser Pro
1090 1095 1100

Gly Ser Pro Ser Ala Pro Asn Arg Pro Gln Gln Ala Asp Gly Ser Pro
1095 1110 1115 1120

Asn Gly Ser Thr Ala Glu Glu Gly Gly Phe Ala Trp Asp Asp Asp
1125 1130 1135

Phe Pro Leu Met Thr Ala Lys Ala Ala Phe Ala Met Ala Leu Asp Pro
1140 1145 1150

Ala Ala Pro Ala Pro Ala Ala Pro Thr Pro Thr Pro Ala Pro Phe Ser
1155 1160 1165

Arg Phe Thr Val Ser Pro Ala Pro Thr Ser Arg Phe Ser Ile Thr His
1170 1175 1180

Val Ser Asp Ser Asp Ala Glu Ser Lys Arg Gly Pro Glu Ala Gly Ala
185 1190 1195 1200

Gly Gly Glu Ser Lys Glu Ala
1205

<210> 17

<211> 1803

<212> DNA

<213> Pan troglodytes

<400> 17

gctccctgcc tggttacacc ctccctggaca gagacagccg gtagtggggg tgaccacccg 60

cagggcagagc ccaagcttgc cacggaggct gagggcactg ccggaccctg tctgcccctt 120

ccttccgtcc cctccccatc ccaggaggga gccccacttc cctcggagga ggccagtgcc 180

cctgacgccc ctgatgccct gcctgactct cccatgcctg ctactggtg cgaggtgtct 240

gccatcaagc tggcttctgt cctgaatggc agcagcagct ctcccgaggt ggaggcaccc 300

agcagcggagg atgaggacac ggctgaggcc acctcaggca tttcaccga cacgtccagc 360

gacggccctgc aggccgagag gctggatgtg gtgccagcct tccgctctct gcagaaggcag 420

gtggggaccc ccgactccct ggactccctg gacatcccat cctcagccag tgatggtggc 480

tatgaggtct tcagccgtc gcccactggc ccctctggag ggcagcccg agcgctggac 540

agtggctatg acaccgagaa ctatgagtcc cctgagttt tgctcaagga ggccgcaggaa 600

gggtgtgagc cccaggcctt tgaggagctg gcctcagagg gtgagggccc cggccccggg 660

Y. Zhou et al.

cccgagacgc ggctctcac ctcccagt ggcctaacs agaagaatcc ctaccgagac 720
tctgcctact tctcagacct ggaggctgag gccgaggccg aggccaccc aggcccagag 780
aagaagtgcg gcggggacca agccccggg ccagagctgg acctgccgag cactggcag 840
ccgtctgagc aggtctccct caggcctggg gttccgggg aggcacaagg ctctggcccc 900
ggggaggtgc tgccccact gtcggctt gaaggatcct cccagagcc cagcacctgc 960
ccctcgggcc tggtccaga gcctccggag cccaaaggcc cagccgaggt gcggcctggg 1020
cccagccca gctgctccca gttttcctg ctgaccccg ttccgctgag atcagaaggc 1080
aacagctctg agttccaggg gccccagga ctgttgcag ggccggcccc aaaaaagcgg 1140
atggggggcc taggcacccc cagagccca ctccgcctgg ctctgcccgg cctccctgcg 1200
gccttggagg gcccggcga ggaggaggag gaggacagtg aggacagcgg cgagtctgac 1260
gaggagctcc gtcgtacag cgtccaggag cctagcgagg acagcgaaga ggaggcgccg 1320
gcggtgcccg tggtggtggc tgagagccag agcgcgcga acctgcgcag cctgctcaag 1380
atgcccagcc tgctgtccga ggcattctgc gaggacctgg aacgcaagaa gaaggccgtg 1440
tccttcttcg acgacgtcac cgtctaccc tttgaccagg aaagccccac ctgggagctc 1500
ggggagccct tcccggcgc caaggaatcg ccccccacgt tccttagggg gagccccggc 1560
tctcccagcg ccccaaccg gccgcagcag gctgatggct cccaaatgg ctccacagcg 1620
gaagagggtg gtgggttcgc gtgggacgac gacttccgc tgatgccggc caaggcagcc 1680
ttcgccatgg ccctagaccc ggccgcaccc gccccggctg cggccacgccc cgctcccttc 1740
tcgcgttca cggtgtcgcc cgccgcacg tccacgtccc gttctccat cacgcacgtg 1800
tct 1803

<210> 18
<211> 1785
<212> DNA
<213> Gorilla gorilla

<400> 18
gctccctgcc tggttacacc ctcctggaca gagacagacg gtagtgggg tgaccacccg 60

caggcagagc ccaagcttgc cacggaggct gagggcactg ccggaccccg cctgcccctt 120
ccttcgtcc cctccccatc ccaggaggga gccccacttc cctcgagga ggccagtgcc 180
cccgacgccc ctgatccct gcctgactcg cccacgcctg ctactggtgg cgaggtgtct 240
gccaccaagc tggctccgc cctgaatggc agcagcagct ctcccgaggt ggagggcaccc 300
agcagtgagg atgaggacac ggctgaggca acctcaggca tcttcaccga cacgtccagc 360
gacggcctgc aggccgagag gcaggatgtg gtgccagcct tccactctct gcagaagcag 420
gtggggaccc ccgactccct ggactccctg gacatccctg cctcagccag tcatggtggc 480
tatgaggtct tcagccgtc ggccacgggc ccctctggag ggcagcccg agcgctggac 540
agtggctatg acaccgagaa ctatgagtcc cctgagtttg tgctcaagga ggcgcaggaa 600
gggtgtgagc cccaggcctt tgcggagctg gcctcagagg gcgagggccc cgggcccag 660
acgcggctct ccacccctt cagtcggctc aacgagaaga atccctaccg agattctgcc 720
tacttcttag acctggaggc tgaggccag gctacccctg gccagagaa gaagtgcgg 780
ggggaccaag cccccgggccc agagctgggc ctgcccagca ctgggcagcc gtctgagcag 840
gtctccctca gtccctgggt ttccgtggag gcacaaggct ctggcccccgg ggaggtgctg 900
cccccaactgc tgcggcttga agggtcctcc ccagagccca gcacctgccc ctgggcctg 960
gtcccagagc ctccggagcc ccaaggccca gccgaggtgc ggcctgggccc cagccccagc 1020
tgctccctgt ttttcctgct gaccccggtt ccgctgagat cagaaggcaa cagctctgag 1080
ttcaggggc ccccaggact gttgtcaggg ccggcccccac aaaagcggat ggggggccc 1140
ggcaccccca gagccccaca ccgcctggct ctgccccggcc tccctgcggc cttggagggc 1200
cggccggagg aggaggagga ggacagttag gacagcgacg agtctgacga ggagctccgc 1260
tgctacagcg tccaggagcc tagcgaggac agcgaagagg aggcgcggc ggtgccctg 1320
gtggtggctg agagccagag cgcgcgcaac ctgcgcagcc tgctcaagat gcccagcctg 1380
ctgtccgagg cttctgcga ggacctggaa cgcaagaaga aggccgtgtc cttcttcgac 1440
gacgtcaccg tctacccctt tgaccaggaa agccccaccc gggagctcgg ggagcccttc 1500

ccgggcgcca aggaatcgcc ccccacgttc cttaggggga gccccggctc ttccagcgcc 1560
cccaaccggc cgcatcgaggc tgcgtggctcc ccaaattggct ccacagcgga agagggtgg 1620
gggttcgcgt gggacgacga cttcccgctg atgccggcca aggacgcctt cgccatggcc 1680
ctagacccgg ccgcacccgc cccggctgctg cccacgccccg ctcccttctc ggcgttcacg 1740
gtgtcgcccg cgcccacgta ccgtttctcc atcacgcacg tgtct 1785

<210> 19
<211> 24
<212> DNA
<213> Pan troglodytes

<400> 19
ggtgaggggcc ccggccccgg gccc 24

<210> 20
<211> 18
<212> DNA
<213> Homo sapiens

<400> 20
ggtgaggggcc ccggggccc 18

<210> 21
<211> 18
<212> DNA
<213> Gorilla gorilla

<400> 21
ggcgaggggcc ccggggccc 18

<210> 22
<211> 24
<212> DNA
<213> Pan troglodytes

<400> 22
ctggaggctg aggcccgaggc cgag 24

<210> 23
<211> 18
<212> DNA
<213> Homo sapiens

<400> 23
ctcgaggctg aggccgag 18

<210> 24
<211> 18
<212> DNA
<213> Gorilla gorilla

<400> 24
ctggaggctg aggccgag 18

<210> 25
<211> 18
<212> DNA
<213> Pan troglodytes

<400> 25
cccacgcccc ctccttc 18

<210> 26
<211> 24
<212> DNA
<213> Homo sapiens

<400> 26
cccacgcccc cgccccgtcc cttc 24

<210> 27
<211> 18
<212> DNA
<213> Gorilla gorilla

<400> 27
cccacgcccc ctccttc 18

<210> 28
<211> 24
<212> DNA

<213> Pan troglodytes

<400> 28

cccacgtcca cgtcccgctt ctcc

24

<210> 29

<211> 18

<212> DNA

<213> Homo sapiens

<400> 29

cccacgtccc gcttctcc

18

<210> 30

<211> 18

<212> DNA

<213> Gorilla gorilla

<400> 30

cccacgtccc gcttctcc

18